Automated Parcel Lockers location problem: a numerical experiment for Turin's urban area

Ottaviani Filippo Maria*, Zenezini Giovanni*, De Marco Alberto*

* *Dipartimento di Ingegneria Gestionale e della Produzione, Politecnico di Torino, Corso Duca degli Abruzzi, 24 10129 – Torino – Italy (filippo.ottaviani@polito.it, giovanni.zenezini@polito.it, alberto.demarco@polito.it)*

Abstract: The significant and constant spike in e-commerce purchases experienced in recent years has generated more pressure on Logistics Service Providers' urban activities, with negative consequences on costs and negative externalities (i.e. environmental impacts and traffic congestion). In this context, home deliveries are some of the major sources of inefficiencies in last-mile delivery systems, due to several reasons such as high order frequency, disperse demand, and non-negligible probability of missed delivery. Consolidating final users' demand in Automated Parcel Lockers (APL) represents an effective solution to the operational strains entrenched in last-mile deliveries. However, this solution requires significant investment by private operators and compels space organizers such as public authorities to grant the usage of portions of public space wherein APLs are installed. An accurate ex-ante appraisal of the variables involved is thus needed in order to discern the overall impact of APL on the urban setting. Hence, this study proposes a solution to the Parcel Lockers network design problem. In particular, the aim of the approach is to minimize the lost demand percentage and the number of APLs installed. The dimensions of the APLs are also assessed by the proposed algorithm in a second step. Input variables and parameters of the model are identified through an online survey submitted to final users located in the city of Turin, Italy. Case-scenarios are simulated via 1000 random demand points according to the real population distribution, potentially covered by 33 APL locations located inside the 90sqkm urban area. The purpose of this work is to draw implications for last-mile private actors, urban space planners and policy makers.

Keywords: Last-mile delivery, Parcel Locker, Optimization, Survey.

1. Introduction

Driven by e-commerce last-mile logistics has increased at constant rate in recent years, showing projected compounded annual growth rates of 15% from 2017 to 2023[1]. Thus, Logistics Service Providers have to increase their productivity rates in order to fulfill the increasing request for home deliveries, which are characterized by high frequency of orders, disperse demand and non-negligible probability of a missed delivery (Cagliano *et al.*, 2017). Each home delivery is also time consuming, since drivers need to decelerate, look for parking, unload the parcel and wait for the receiver (Akeb, Moncef and Durand, 2018). Therefore, last-mile home deliveries generate a significant number of vehicle-kilometers for each delivery route, with negative consequences on either the profitability of private operators and the emissions of pollutant (e.g. CO_2, NO_x, PM10) (Schliwa *et al.*, 2015). Consolidating final users' demand in Automated Parcel Lockers (APL) represents an effective solution to the operational strains entrenched in last-mile deliveries, by virtue of reducing the points of delivery and consequently the average delivery time (Bailey *et al.*, 2013). Parcel lockers are a type of unattended collection-and-delivery point (CDP) installed in public and private areas, where parcels are retained for a limited amount of time until the customer is able to retrieve them by using the order reference code (Iwan, Kijewska and Lemke, 2016). However, this solution requires significant investment by private operators and compels space organizers such as public authorities to grant the usage of portions of public space wherein APLs are installed (Zenezini *et al.*, 2018). Moreover, the diffusion of parcel lockers depends on the potential customers' perception and inclination towards the adoption of technological innovations that impose a change of customers behavior. Vakulenko, Hellström and Hjort (2018) found that co-value in last-mile logistics is generated between parcel lockers operators and customers, but that this process of value co-generation can also lead to reduction of value when delivery issues occur. Besides the ease of use of this new technology, that might crowd-out less tech-savvy customers, one major variable driving the adoption of parcel lockers is location. Traditionally, parcel lockers are installed in easily accessible places, controlled and close to places with a high frequency of shipments (service stations, shopping malls, squares) (Janjevic, Kaminsky and Ndiaye, 2013). Using this type of facilities lowers the entry barriers because they provide a solution to safety issues often encountered in customers surveys (Lachapelle *et al.*, 2018) and help maximizing the catchment

[1] https://www.emarketer.com/content/global-ecommerce-2019\

area of parcel lockers. Moreover, these facilities are mostly reached by car by customers on the route to/from their homes, and thus the nuisance of picking up parcels is reduced. Given the undeniable operational benefits of such delivery option and factoring in the estimated continuous increase in online purchases, we might expect more widespread adoption of APLs in cities worldwide. Preferred locations nowadays might not be enough to cover all the future demand. Furthermore, understanding the implications on quantifying the number of APL locations of customers' willingness to retrieve the parcels by themselves. An accurate ex-ante appraisal of the variables involved is thus needed to discern the overall impact of APL on the urban territory. Moreover, existing papers do not provide a comprehensive methodological framework able to guide practitioners through the different steps of the problem, from data collection to visualizing the APL stations on the territory. To this end, this study proposes a version of the location set covering problem to the Parcel Lockers network design problem, which considers not only the distances that customers are potentially willing to cover in order to retrieve the parcel but also the demand lost as a consequence of not reaching the customer through an APL location. The purpose of this work is to draw implications for last-mile private actors and space planners.

The paper is structured as follows. First, a review of pertinent literature is proposed. Then, in Section 3 the research methodology is outlined in detail, with the results of a numerical experiment presented in Section 4. Finally, discussions and conclusions are drawn in Section 5.

2. Literature Review

The APL location problem represents a niche yet emerging stream of research within the domain of last-mile logistics.

Lachapelle *et al.* (2018) provide a comprehensive assessment of existing APL locations in terms of customers' preferred attributes such as safety, proximity to highways and accessibility of the location. Thus, the authors draw implications for future locations based on existing ones, which might be the result of faulty evaluation in the first place, rather than through an optimization algorithm. Mathematical formulations aimed at optimizing the number and location of facilities within a network of nodes fall under the generic term of Facility Location problems. Most of such problems are p-median problems, location set covering problems and location-allocation problems.

In *p-median* problems, the objective function minimize the total demand-weighted distance between customers and facilities. Kedia, Kusumastuti and Nicholson (2019) adopt a consumer-centric approach to Collection-and-delivery points location, submitting a survey to users in order to retrieve the demand parameters and maximum distance covered by consumers willing to pick up the parcel at a CDP location. Deutsch and Golany (2018) propose an un-capacitated facility location problem. In their model, demand is lost if customers are not reached by at least one APL and their willingness to move decreases with the distance to the parcel locker. Location set covering problems minimize the cost of installing P facilities needed to cover a specified level of demand. (Lee *et al.*, 2019) adopt a two-step approach to the APL location problem. First, they identify the set of potential locations according to the afore-mentioned requirements. Second, they use a set-covering model to estimate which ones of the potential locations should receive an APL. However, the decision criteria used to select the potential locations are unclear and the numerical case study comprises a relatively small neighbourhood. Moreover, set covering problems generally risk generating too large numbers of facilities because nodes need to be covered regardless of their individual demand.

Previous works show that APL location problem can be tackled with a long-term planning horizon. More recent studies look at the impact of APL locations on a short-term planning horizon, integrating the use of such delivery solution into the Vehicle-Routing Problem (VRP), therefore adopting a location-allocation approach. Orenstein, Raviv and Sadan (2019) for instance optimize the assignment of parcels to both the vehicles and APL modules, but assumes equal attractiveness of different APL locations for the customers, thus not considering any customers' preference in the optimization algorithm. Enthoven *et al.* (2020) instead aim to minimize costs for both customers and LSPs and include penalties if customers are not reached via their preferred delivery method.

Hence, the main objective of an APL location problem is to find the location that minimizes the number of APLs to install and operational costs considering final customers in the decision. These models consider a fixed distance allowed between the APL and the customer points. Final customers' however do not all share the same level of appreciation and commitment for this last-mile solution, and thus a location model should consider a certain amount of variability in customers' willingness to move from their place of residence to pick up the parcel as well as a variability in the demand of parcels. Because of their preferences and the multiple delivery options that customers can now choose from, failing to cover a customer from an APL location will almost certainly result in lost demand. Finally, our model does not seek to optimize vehicle routing but rather to identify a set of locations to cover the potential demand by customers. Hence, we propose a location set covering problem which aims at overcoming the issue of choosing too many facilities by weighting each node (i.e. customer) with its demand.

3. Methodology

3.1 Model

Since we have used MS Excel software both to setup the problem and run the optimizations, the model constraints are described both using logical statements to reflect the formulas entered in the spreadsheet cells, and mathematical terms.

The problem we are facing can be considered as a subgenre of the Facility Location Problems, its aim is to limit the impact of the facilities to install inside the city urban system while satisfying a prefixed amount of estimated demand. The parameters included in the model are generated according to the statistics obtained by the Survey we made.

The main assumption undertaken during the problem modelling is that the APLs have been considered to have got unlimited volumetric capacity, implying they can satisfy any amount of demand by simply being located at a distance inferior to the satisfied users' tolerances.

$$\text{minimize} \sum_{j=1}^{J} active_j \tag{1}$$

$$dem\_lost\_pct \leq \text{dem\_lost\_pct\_max} \tag{2}$$

$$dem\_lost\_pct = dem\_lost/\text{dem\_tot} \tag{3}$$

$$\text{dem\_tot} = \sum_{i=1}^{I} \text{dem}_i \tag{4}$$

$$dem\_lost = \sum_{i=1}^{I} (1 - satisfied_i) \cdot \text{dem}_i \tag{5}$$

$$\text{IF tol}_i \geq min\left(d_{ij}\right) \text{ THEN } satisfied_i = 1 \tag{6}$$

$$\text{IF } active_j = 1 \text{ THEN } d_{ij} = \sqrt{\left(\mathrm{x}_i - x_j\right)^2 + \left(\mathrm{y}_i - y_j\right)^2} \text{ ELSE } d_{ij} = M \tag{7}$$

$$\min(\text{lower\_x}_i) \leq x_j \leq \max(\text{upper\_x}_i) \tag{8}$$

$$\min(\text{lower\_y}_i) \leq y_j \leq \max(\text{upper\_y}_i) \tag{9}$$

$$active_j \in \{0,1\} \tag{10}$$

$$satisfied_i \in \{0,1\} \tag{11}$$

$$\text{dem\_lost\_pct\_max} \in \mathbb{R}^+ \tag{12}$$

$$\mathrm{x}_i, \mathrm{y}_i, x_j, y_j \in \mathbb{Z} \tag{13}$$

$$I = \{1..1000\} \tag{14}$$

$$J = \{1..33\} \tag{15}$$

The objective function (1) is to minimize the number of APLs to locate, variable (10) being Boolean describes the activation or not of the single APL and the maximum value the objective function can reach is given by parameter (15).

The main constraint (2) the model is subject to is that the total unsatisfied demand percentage must not exceed a prefixed amount given by parameter (12). The total unsatisfied demand percentage (3) is given by the ratio between the total unsatisfied demand and the total estimated demand. The total estimated demand (4) is obtained by summing all the individual users' demands. The total unsatisfied demand (5) is obtained by summing the individual users' demand multiplied by their satisfaction factor (11).

Statements (6) and (11) define the Boolean variable and force its value to 1 whenever the user's tolerated distance to travel is equal or exceeds the air distance between the coordinates the user is set in and the nearest APL facility, otherwise 0. Variable (11) being Boolean implies we are assuming a user's demand can be either completely satisfied or not. Both the users' randomized locations, as problem parameters, and the possible APLs locations, as problem variables, are identified using the two-dimensional Cartesian *x,y* coordinate system. The air distance between the user and the APLs is given by statement (7) that can be interpreted as follows: whenever variable (10) equals 0 the APL is not activated and all the distances between that very APL and all the users is set to big M so none of them, together with constraint (6), will be considered to be satisfied.

3.2 Parameters

Before being considered as static values, the model parameters have been established beforehand.

The maximum number of possible APLs to locate (15) has been capped to the number of different postcodes the urban city of Turin has got. This choice is coherent to the fact postcodes are used by the national postal service of Italy to encode the addresses of correspondence, also it allows the facilitation of the sorting mails work and copes with the huge growth of postal traffic.

The number of users to generate (14) has been capped to represent just the 0.1% circa of the actual urban population. Since both the problem setup and optimization have been carried out using Microsoft Excel software, instead of setting the users sample dimension to the actual population of 879'004 (circa) inhabitants and run the optimization just once, we had to cap the number to a reasonable amount (14). The choice of the number 1000 ensures that randomized data for the users' parameters satisfy the significance tests needed to verify their values fit the theorized distributions they have been generated from.

The maximum percentage of unsatisfied demand (12) allowed must be chosen by the model developers.

Each user's individual demand and tolerance have been randomized from an estimated distribution. The data to evaluate the distributions from has been collected through a survey conducted in the city of Turin to e-commerce users. Through the survey, which received 655 useful responses, the parameters for the exponential distributions for both the parcels demand and customers' tolerance in meters were identified. The distributions have been previously analysed to verify the absence of correlation and causation relationships between the different parameters involved inside the model. The demand generated is shown in Equation 16.

$$dem_i \sim \exp(37.2)$$
$$dem_i = \frac{\ln(1 - d\_rand()_i)}{-1/37.2} \tag{16}$$

The sub-parameter (17) is generated through the uniform distribution between the values 0 and 1

$$d\_rand()_i \sim \mathrm{U}(0,1) \tag{17}$$

The tolerance parameter (18) is expressed in metres and has been randomized from the estimated distribution.

$$tol_i \sim 6670 \cdot \beta(1.23, 3.27)$$
$$tol_i = 6670 \cdot \beta^{-1}(t\_rand()_i, 1.23, 3.27) \tag{18}$$

The sub-parameter (19) is generated through the uniform distribution between the values 0 and 1.

$$d\_rand()_i \sim \mathrm{U}(0,1) \qquad (19)$$

We have evaluated the correlation Pearson product-moment correlation coefficient between each parameter and verified its significance through its *t*-test (20): all variables resulted not correlated.

$$\begin{gathered}
\rho_{a,b} = \frac{\sum_{k=1}^{K}(a-\bar{a})(b-\bar{b})}{\sqrt{\sum_{k=1}^{K}(a-\bar{a})^2 \sum_{k=1}^{K}\left(b-\bar{b}\right)^2}} \\
H_0 : \rho_{a,b} = 0 \\
H_1 : \rho_{a,b} \neq 0 \\
t = \rho_{a,b}\sqrt{\frac{n-2}{1-\rho_{a,b}}} \\
\text{IF } t > t_{1-\frac{p}{2}} \text{ THEN } H_1 \text{ ELSE } H_0
\end{gathered} \qquad (20)$$

The estimated distributions used to randomize the individual demand and tolerance parameters have been evaluated using Rockwell's Arena Simulation Input Analyzer software: the criterium adopted to choose the best distribution to fit the data is the minimum square error criterium, for both the two formerly cited parameters the corresponding *p*-values are such that both χ^2 and Kolmogorov-Smirnov Tests are satisfied.

The users' *x,y* coordinates are randomly generated through uniform distributions (21) between the respective lower and upper bounds.

$$\begin{gathered}
x_i \sim \mathrm{U}(\mathrm{lower_x}_i, \mathrm{upper_x}_i) \\
y_i \sim \mathrm{U}(\mathrm{lower_y}_i, \mathrm{upper_y}_i)
\end{gathered} \qquad (21)$$

The values the bounds (22) assume depend on the interval the sub-parameter (23) belongs to.

$$(\mathrm{lower_x}_i, \mathrm{upper_x}_i, \mathrm{lower_y}_i, \mathrm{upper_y}_i) = \begin{cases}
-1500, 500, -2200, 1600 & \mathrm{circ_rand}()_i \leq 0.09 \\
-4200, -500, -6300, -1800 & 0.09 < \mathrm{circ_rand}()_i \leq 0.245 \\
-6000, -1600, -2300, 500 & 0.245 < \mathrm{circ_rand}()_i \leq 0.386 \\
-6000, -1600, -500, 1600 & 0.386 < \mathrm{circ_rand}()_i \leq 0.496 \\
-6000, -1600, 2500, 5200 & 0.496 < \mathrm{circ_rand}()_i \leq 0.637 \\
-1600, 1600, 2800, 5000 & 0.637 < \mathrm{circ_rand}()_i \leq 0.758 \\
-1600, 3000, 1600, 3000 & 0.758 < \mathrm{circ_rand}()_i \leq 0.855 \\
-1100, 2200, -5300, 1600 & > 0.855
\end{cases} \qquad (22)$$

The intervals bounds are linked to the cumulative distribution of the real Turin population among the 8 districts the urban area is divided into. Reference data have been collected from the Municipality statistics office.

The process of identifying the districts and postcodes bounds has been realized using the Google Maps Measure distance feature, the point corresponding to coordinates (0,0) located in Turin Porta Nuova metro station.

The sub-parameter (23) value is generated through a uniform distribution between the values 0 and 1.

$$circ\_rand()_i \sim \mathrm{U}(0,1) \qquad (23)$$

3.3 Algorithm Initialisation

The optimization has been carried out with the Microsoft Excel Solver add-in, developed by Frontline Solvers ®.

The solving method we have selected is the Evolutionary one and its parameters have been set as follows:

Table 1: Solver evolutionary method settings

Constraint Precision = 0.000001
Use Automatic Scaling = TRUE
Mutation rate = 0.1
Population size = 98
Random seed = 0
Maximum number of subproblems = 500

Since the method chosen relies on genetic and evolutionary algorithms, to speed up the calculations both variables initialization and bounds identification have been carried out.

The APLs coordinates have been forced to be integer values (13) and limited to the area the randomized population is generated from (8, 9), initialized in the centre of gravity of each one of the 33 postcodes CAP, equal to the average of the coordinates bounds (24).

$$\begin{gathered}
x_j = (\min(\mathrm{x}_i) + \max(\mathrm{x}_i))/2 \\
y_j = (\min(\mathrm{y}_i) + \max(\mathrm{y}_i))/2
\end{gathered} \qquad \forall j \in J \qquad (24)$$

then de-activated it (25).

$$active_j = 0 \qquad \forall j \in J \qquad (25)$$

4. Numerical Experiment

Instead of fixing the parameter (12) and randomizing the users sample many times to check how much the variable (3) changes once an optimization is done and its solution fixed, our intent is to understand for which value of the variable (12) the number of APLs starts converging to its cap (15).

The initial baseline scenario assumes an initial value of unsatisfied demand allowed of 35%. Figure 1 shows the solution for the algorithm on the map of Turin. We then proceed on reducing the allowed unsatisfied demand by 5% at a time. Figure 2 shows the number of APLs according to different unsatisfied demand parameters. It is clear from the figure that the number of locations needed ramps up below a 30% threshold, meaning that it is more difficult to cover marginal demand.

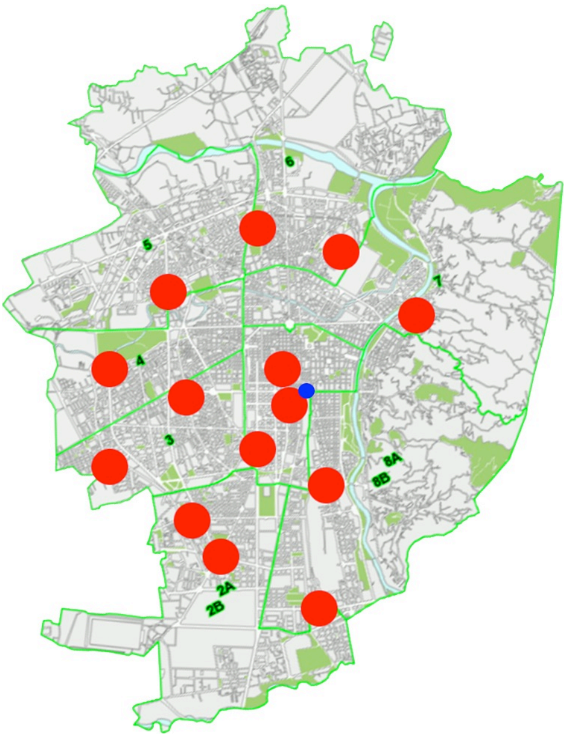

Figure 1 Location and number of APLs for an unsatisfied demand of 30%

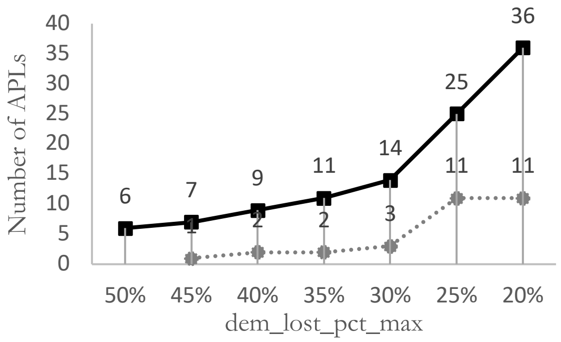

Figure 2 Number of APL to activate according to different unsatisfied demand thresholds

The identified solutions may have a limitation insofar as the algorithm may not reach the optimal solution in due time. However, by randomizing the sample of users and running the optimizations again, the values evaluated in (12) would still be the same (not considering extreme cases or exceptions).

Another implicit assumption of the model is that the variable (11) doesn't take into consideration the variance of each user's tolerance neither it allows the demand to be accounted for just a percentage, possibly according to how the distances between the users and the nearest APLs is bigger than their tolerances.

5. Discussion and Conclusion

Our study intended to describe a framework to approach the last-mile issue concerning APLs identification and location.

To optimize the problem, we have used Microsoft Excel Solver where the different model constraints are represented by the very formulas evaluated inside the spreadsheet cells. Model parameter involved both subjective parameters, such as the willingness for customers to move to the parcel locker and customers' demand, and objective parameters as population distribution across the considered geographical area. Subjective parameters were retrieved from a survey to e-commerce users.

The optimization algorithm was tested with a numerical experiment based on the city of Turin. Since it was possible to assume that the urban population is homogenous towards APL utilization, the optimized solution scattered the different identified APLs across the considered area, placing more than one facility in the most populated ones. The number of the APLs to locate, instead, seems to follow an exponential distribution converging to the half of the maximum urban postcodes (14 ~ 16 versus 33) for a reasonable amount of maximum estimated percentage of unsatisfied demand equal to 30%.

Through this study, space planners are able to discern the number and potential locations to allow to private operators for setting up an APL network. LSPs on the other hand are able to have an estimation of how many APLs are needed to cover a predetermined amount of demand by the final users.

This study has some limitations. Difficulties and assumptions play a very important role in the problem modelling: all the model parameters rely on estimated distributions and maths linked to human behaviour which could change drastically from time to time and from one person to the other one; the optimization being carried out takes a not negligible amount of time before converging to a solution, and this might be a problem when increasing the number of users to cover the whole population of the city.

Future research will focus on adding new constraints to the model, such as the volumetric capacity of the APLs or the minimum distance between two facilities, and to change the optimization method to a mathematical one, as convex quadratic mixed integer modelling models, to both shorten the optimization times and provide a more reliable global solution.

References

Akeb, H., Moncef, B. and Durand, B. (2018) 'Building a collaborative solution in dense urban city settings to enhance parcel delivery: An effective crowd model in Paris', *Transportation Research Part E: Logistics and Transportation Review*, 119, pp. 223–233. doi: 10.1016/j.tre.2018.04.007.

Bailey, G. *et al.* (2013) 'Can locker box logistics enable more human-centric medical supply chains?', *International Journal of Logistics Research and Applications*. Taylor & Francis, 16(6), pp. 447–460.

Cagliano, A. C. *et al.* (2017) 'Levers of logistics service providers' efficiency in urban distribution', *Operations Management Research*, 10(3–4). doi: 10.1007/s12063-017-

0125-4.

Deutsch, Y. and Golany, B. (2018) 'A parcel locker network as a solution to the logistics last mile problem', *International Journal of Production Research*. Taylor & Francis, 56(1–2), pp. 251–261.

Enthoven, D. L. J. U. *et al.* (2020) 'The Two-Echelon Vehicle Routing Problem with Covering Options: City logistics with cargo bikes and parcel lockers', *Computers & Operations Research*. Elsevier, p. 104919.

Iwan, S., Kijewska, K. and Lemke, J. (2016) 'Analysis of Parcel Lockers' Efficiency as the Last Mile Delivery Solution - The Results of the Research in Poland', in *Transportation Research Procedia*, pp. 644–655. doi: 10.1016/j.trpro.2016.02.018.

Janjevic, M., Kaminsky, P. and Ndiaye, A. B. (2013) 'Downscaling the consolidation of goods-state of the art and transferability of micro-consolidation initiatives', *European Transport - Trasporti Europei*, (54).

Kedia, A., Kusumastuti, D. and Nicholson, A. (2019) 'Establishing Collection and Delivery Points to Encourage the Use of Active Transport: A Case Study in New Zealand Using a Consumer-Centric Approach', *Sustainability*. Multidisciplinary Digital Publishing Institute, 11(22), p. 6255.

Lachapelle, U. *et al.* (2018) 'Parcel locker systems in a car dominant city: Location, characterisation and potential impacts on city planning and consumer travel access', *Journal of Transport Geography*. Elsevier, 71, pp. 1–14.

Lee, H. *et al.* (2019) 'Development of a Decision Making System for Installing Unmanned Parcel Lockers: Focusing on Residential Complexes in Korea', *KSCE Journal of Civil Engineering*. Springer, 23(6), pp. 2713–2722.

Orenstein, I., Raviv, T. and Sadan, E. (2019) 'Flexible parcel delivery to automated parcel lockers: models, solution methods and analysis', *EURO Journal on Transportation and Logistics*. Springer, 8(5), pp. 683–711.

Schliwa, G. *et al.* (2015) 'Sustainable city logistics — Making cargo cycles viable for urban freight transport', *Research in Transportation Business & Management*, 15, pp. 50–57. doi: 10.1016/j.rtbm.2015.02.001.

Vakulenko, Y., Hellström, D. and Hjort, K. (2018) 'What's in the parcel locker? Exploring customer value in e-commerce last mile delivery', *Journal of Business Research*, 88, pp. 421–427. doi: 10.1016/j.jbusres.2017.11.033.

Zenezini, G. *et al.* (2018) 'The Collection-And-Delivery Points Implementation Process from the Courier, Express and Parcel Operator's Perspective', *IFAC-PapersOnLine*. Elsevier, 51(11), pp. 594–599.